

法政大学学術機関リポジトリ  
HOSEI UNIVERSITY REPOSITORY

# 主要信号抽出とウェーブレット変換による非接触脈 拍数推定法

著者	坂巻 燎
出版者	法政大学大学院情報科学研究科
雑誌名	法政大学大学院紀要．情報科学研究科編
巻	16
ページ	1-6
発行年	2021-03-24
URL	<a href="http://doi.org/10.15002/00023869">http://doi.org/10.15002/00023869</a>

# 主要信号抽出とウェーブレット変換による非接触脈拍数推定法

## Remote Pulse Rate Estimation Method Based on Principal Signal Extraction and Wavelet Transform

坂巻 燎\*

Ryo Sakamaki

法政大学大学院 情報科学研究科 情報科学専攻

Email: 19t0008@cis.k.hosei.ac.jp

**Abstract**—In this paper, we propose a remote pulse rate estimation method from video under fluorescent light. Our method focuses on the following two assumptions. One is that the pulse rate is approximately equivalent at any points on the skin. Another is that the pulse rate does not change so quickly while keeping us at rest. First of all, we track face to extract skin region from video. Next, we divide the skin region into blocks and convert the time sequence of color images within the region to color signals. After that, we de-trend each color signal and convert them to chrominance signals. Then, we extract principal signal from chrominance signals by pruning of signal with different peak intervals and using singular value decomposition, under the assumption of the pulse rate equality over any surface of the skin. Finally, we use wavelet transformation to estimate the pulse rate under the assumption of the stability of the pulse rate while resting. As a result, our method achieves improving estimation accuracy. The mean absolute error is 0.762bpm, the mean absolute percentage error is 1.04%, the standard deviation of error is 0.899bpm, the correlation coefficient is 0.987 and, in 90% of data, the error is less than 2bpm.

### 1. 研究背景と目的

2000年代に入り、世界各国では65歳以上の高齢者人口が増加している。特に、日本やイタリアなど一部先進国では総人口に対する65歳以上人口の比率は2018年の時点で20%を超え、高齢化問題が顕在化している。そのため、地方に住む高齢者の見守りサービスの開発の必要性が高まってきている。また、2020年に発生した新型コロナウイルス(SARS-CoV-2)の世界的流行によって、個人の健康への意識や遠隔診断の必要性の増加、感染防止対策という観点からも日々の健康状態やバイタルサインを遠隔からチェックできるシステムの重要性は高まると考えられる。健康状態をチェックするシステムの例としては脈波計や血圧計といった接触型の機器のほか、高齢者を遠隔から見守るサービスとして、センサーマットやカメラなどを使い異常を検知するといったシステムがある。しかし、接触型の機器は計測準備や接触感染防止のための消毒の手間がかかり、高齢者の見守りサービスでは設置のコストがかかる。

これらに対し、比較的簡単に脈拍数を得られる方法としてリモート光電式容積脈波記録法(remote Photo-plethysmography; rPPG)がある。この方法では、自然光や蛍光灯の光の下で、測定する人物の顔の動画をスマートフォンなどのカメラで撮影し、顔から反射した光を調べることで被写体の脈拍数を推定する。反射光から脈拍数を推定する方法としては主に、酸素化ヘモ

グロビンの吸光特性を利用し、緑色成分に対してフーリエ変換を適用する方法[1]や独立成分分析を用いて統計的に独立な信号へ分離し、分離された信号から推定する方法[2]などが用いられている。しかし、脈波の振幅よりも身じろぎなどによって発生する波の振幅の方が大きいため、血圧計などで測った脈拍数の実測値と大きく異なることがあった。

以上の研究状況のもと、本研究では皮膚のどの箇所を調べても脈拍数は等しいことと、脈拍数に対応する周波数は急激には変化せず長期にわたって観測されることに着目し、主要な信号の抽出とウェーブレット変換を用いた脈拍数推定方法について提案する。そして、関連研究で挙げられた推定方法と本研究の方法を実測値との誤差の平均や実測値との相関係数などの基準で比較し、測定精度を検証する。

### 2. 関連研究

rPPGの測定原理は、心臓の拍動によって、血管内の血液量が僅かに変化することに基づく。この原理に基づく、皮膚に光を当てた時の光の吸収、反射の程度も拍動に従って変化する。また、酸素化ヘモグロビンの吸光特性を用いると動画像の緑色成分に脈波が強く現れることが分かっている[1]。この性質を利用して動画像の緑色成分に対し、頬や額といった皮膚の部分領域の領域内平均を毎フレーム求め、得られた信号にフーリエ変換を適用することで脈拍数を推定している[1]。この方法は非常に簡単に求められるものの、身じろぎといった僅かな体の動きにも影響を受けやすいという弱点がある。

独立成分分析(ICA)を用いて脈拍数を求める方法もある[2]。この方法では動画像の赤、緑、青色成分に対し、頬や額といった皮膚の部分領域の領域内平均を毎フレーム求め、得られた信号にICAを用い互いが統計的に独立した信号に分離をする。その後、分離した信号の中からパワースペクトルの最大値が最も強く現れる信号を選択し脈拍数を推定する。

Chrom法[3]はまず式(1)の変換式を用いて赤、緑、青色成分の信号に対し、トレンド除去を行い、光源の影響などを抑えた信号 $S_r^d$ 、 $S_g^d$ 、 $S_b^d$ を求める。そして、式(1)を用いて $S_{chrom}$ という信号を求める。類似した手法で、POS(Plane-Orthogonal-to-Skin)と呼ばれる方法[4]では式(2)中の信号 $S_{pos}$ を用いて推定する。

$$\begin{pmatrix} X_{chrom} \\ Y_{chrom} \end{pmatrix} = \begin{pmatrix} 3 & -2 & 0 \\ 1.5 & 1 & -1.5 \end{pmatrix} \begin{pmatrix} S_r^d \\ S_g^d \\ S_b^d \end{pmatrix}$$
$$\alpha = \frac{\sigma(X_{chrom})}{\sigma(Y_{chrom})} \quad (1)$$
$$S_{chrom} = X_{chrom} - \alpha Y_{chrom}$$

\* Supervisor: Prof. Satoru Fujita

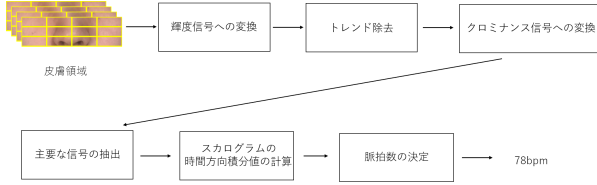


図 1. 脈拍数推定の処理手順

$$\begin{pmatrix} X_{pos} \\ Y_{pos} \end{pmatrix} = \begin{pmatrix} 0 & 1 & -1 \\ -2 & 1 & 1 \end{pmatrix} \begin{pmatrix} S_r^d \\ S_g^d \\ S_b^d \end{pmatrix}$$

$$\alpha = \frac{\sigma(X_{pos})}{\sigma(Y_{pos})} \quad (2)$$

$$S_{pos} = X_{pos} + \alpha Y_{pos}$$

ただし、 $\sigma(\cdot)$  は標準偏差を求める操作を表す。また、2SR(Spatial Subspace Rotation) と呼ばれる方法 [5] では、被写体の肌の色の空間分布を求め推定している。

これらの方法に対して、本研究では皮膚のどの箇所を調べても脈拍数はほぼ等しいことと、脈拍数に対応する周波数は急激には変化せず長期にわたって観測されることに着目し、主要な信号の抽出とウェーブレット変換を用いた脈拍数推定方法について提案する。

### 3. 推定方法

#### 3.1. 皮膚領域の取り出し

撮影した顔の動画像から脈拍数を推定するには初めに顔のトラッキングとその中の皮膚領域を取り出す必要がある。顔の検出には Viola-Jones 法、トラッキングには KLT 法を用いる。そして、取り出す皮膚領域は鼻と両頬を含む矩形領域とし、Algorithm1 によって得られる *skin* である。この方法はまず 3 から 5 行目で検出された顔領域を傾きのない矩形に変換する。これにより、皮膚領域の決定や取り出す操作を簡単にすることができる。次に、7 行目で目や口、髪といった皮膚に動きがある部分や皮膚でない部分を除く。そして、最終的に皮膚領域を決める際、顔領域の中心の色は被写体の皮膚の色であると仮定し、HSV 色空間上でその色と近い色なるべく多く含むように矩形領域を設定している。

皮膚領域として取り出す部分を決めた後はその部分を、継続的に毎フレーム取り出す。この方法は Algorithm1 で得られた *bbox2* と *skin* を用いて Algorithm2 のようにして行う。Algorithm2 で行っていることは、6 行目でホモグラフィ変換を用いて *bbox2* と同じ座標系へ変換し、8 行目で同じ皮膚の部分を取り出している。

この後の得られた皮膚領域から脈拍数を推定する手順を図 1 に示す。まず、得られた皮膚領域から輝度の信号を計算し、得られた信号のトレンド除去を行う。次に式 (1) で得られるクロミナンス信号に変換し、異なるピーク間隔を持つ信号を取り除いた後、特異値分解を用いて主要な信号の抽出を行う。最後に、得られている各信号に対しスカログラムの時間方向の積分値を計算し、脈拍数を決定する。

#### 3.2. 前処理

得られた皮膚領域から輝度の信号を計算する。その際、皮膚領域を  $3 \times 4$  のブロックに分け、RGB 各 3 チャネルについて計算し合計 36 個の輝度の信号を得る。 $i$  行  $j$

#### Algorithm 1 皮膚領域の決定

```

 $I \leftarrow$  1 フレーム目の画像
 $bbox1 \leftarrow$   $I$  の顔を囲む 4 つの頂点
 $bbox2 \leftarrow$  左上  $x, y$  座標を  $bbox1$  の  $x, y$  座標の最小値,
右下  $x, y$  座標を  $bbox1$  の  $x, y$  座標の最大値とする矩形領域の 4 つの頂点
 $H \leftarrow (bbox2 = H \text{ bbox1})$  となるホモグラフィ行列
 $I2 \leftarrow I$  を  $H$  を用いて変換した画像
 $face \leftarrow I2$  から  $bbox2$  の部分を切り取った画像
 $face$  から上 40% (下まぶたより上の部分) と下 20% (上唇より下の部分) を取り除く
 $w \leftarrow face$  の横幅
 $h \leftarrow face$  の縦幅
 $tol \leftarrow 0.1w \times 0.1h$ 
 $sample \leftarrow face$  の中心  $0.1h \times 0.1w$  ピクセル
 $skin \leftarrow face$  の中心  $0.2h \times 0.2w$  ピクセルを含む矩形領域の 4 点
 $h_{mu} \leftarrow sample$  の色相の平均
 $h_{std} \leftarrow sample$  の色相の標準偏差
 $s_{mu} \leftarrow sample$  の彩度の平均
 $s_{std} \leftarrow sample$  の彩度の標準偏差
 $pixs \leftarrow sample$  のピクセル数
while true do
     $img \leftarrow face$  から  $skin$  の部分を切り取った画像
     $cpixs \leftarrow img$  の色相が  $h_{mu} \pm h_{std}$  の範囲に入り,
    彩度が  $s_{mu} \pm s_{std}$  の範囲に入るようなピクセルの総数
     $d \leftarrow cpixs - pixs$ 
    if  $d < 0$  かつ  $|d| > tol$  then
        break
    else if  $|d| < tol$  then
        break
    end if
     $skin$  の縦幅と横幅を 20% 増やす
    if  $skin$  が  $face$  よりも大きい then
         $skin \leftarrow face$  の 4 つの頂点
        break
    end if
end while
 $skin$  の値を  $I2$  の左上が原点になるように補正

```

#### Algorithm 2 皮膚領域の取り出し

```

 $N \leftarrow$  総フレーム数
 $Video \leftarrow$  空のリスト
for frame = 1 to  $N$  do
     $I \leftarrow$  frame 目の画像
     $bbox1 \leftarrow I$  の顔を囲む 4 つの頂点
     $H \leftarrow (bbox2 = H \text{ bbox1})$  となるようなホモグラフィ行列
     $I2 \leftarrow I$  を  $H$  を用いて変換した画像
     $img \leftarrow I2$  から  $skin$  の部分を切り取った画像
     $Video$  の末尾に  $img$  を追加
end for

```

列目のブロックの動画像  $V_{i,j}$  から輝度信号  $S_{i,j}$  への変換は式 (3) で行う。

$$S_{i,j} = [M_{i,j}(1), M_{i,j}(2), \dots, M_{i,j}(t), \dots, M_{i,j}(N)]$$

$$M_{i,j}(t) = \frac{1}{W_{i,j} H_{i,j}} \sum_{x,y} V'_{i,j,x,y}(t) \quad (3)$$

$$V'_{i,j,x,y}(t) = \text{medfilt}(V_{i,j}(t), x, y)$$

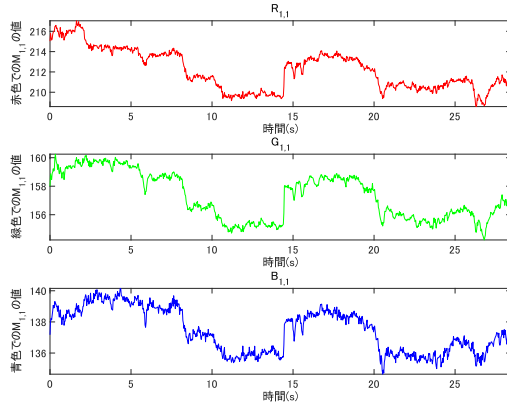


図 2. 1 行 1 列目のブロックの各色成分の輝度信号

ただし,  $N$ ,  $W_{i,j}$ ,  $H_{i,j}$  は動画の総フレーム数,  $V_{i,j}$  の横幅と縦幅を表す. また,  $V_{i,j}(t)$  は  $t$  フレーム目の  $V_{i,j}$  の画像を表す. そして,  $\text{medfilt}(V_{i,j}(t), x, y)$  はフィルタサイズが  $3 \times 3$  のメディアンフィルタを画像  $V_{i,j}$  に適用し, 座標  $(x, y)$  の値を取得する操作,  $V'_{i,j,x,y}(t)$  はその操作によって得られた値を表す. メディアンフィルタを適用する理由は, 皮膚にあるひげや毛穴の影響を取り除くためである. そして, 式 (3) を動画の RGB 各成分に適用し, 得られた輝度信号をそれぞれ  $R_{i,j}$ ,  $G_{i,j}$ ,  $B_{i,j}$  とする.

次に, 得られた信号にトレンド除去を行う. これを行うことで, 時間変化する直流成分や僅かに体が動くことで発生する反射光の変化の影響を取り除くことができる.トレンド除去の方法は文献 [3] で用いられている方法を使用する. すなわち, 対象とする信号  $S$  の移動平均線  $\mu(S)$  を求めた後, 元の信号の各値を移動平均線で割るという方法である.トレンド除去後の信号  $S^d$  は式 (4) で表される.

$$S^d = S \oslash \mu(S) \quad (4)$$

ただし,  $\oslash$  はアダマール除算を表す. そして,トレンド除去後の  $R_{i,j}$ ,  $G_{i,j}$ ,  $B_{i,j}$  をそれぞれ  $R_{i,j}^d$ ,  $G_{i,j}^d$ ,  $B_{i,j}^d$  とする. 図 2, 3 はそれぞれ 1 行 1 列目のブロックに対し, 式 (3) に基づいて作成した赤, 緑, 青色成分の信号, 図 2 の信号を式 (4) に基づいてトレンド除去した信号を表している. また, 図 2 の横軸は秒単位の時間, 縦軸は式 (3) の 2 段目の式を表し, 図 3 の横軸は秒単位の時間, 縦軸はトレンド除去後の値を表している. そして, どちらの図も上から順に赤, 緑, 青色成分の信号を表している.

トレンド除去を行った後はクロミナンス信号に変換する. 変換の方法は式 (1) と同じであるが, それをブロックごとに行う. すなわち, 式 (5) によって, 各ブロックのクロミナンス信号  $S_{i,j}^{chrom}$  を得る.

$$\begin{pmatrix} X_{i,j}^{chrom} \\ Y_{i,j}^{chrom} \end{pmatrix} = \begin{pmatrix} 3 & -2 & 0 \\ 1.5 & 1 & -1.5 \end{pmatrix} \begin{pmatrix} R_{i,j}^d \\ G_{i,j}^d \\ B_{i,j}^d \end{pmatrix}$$

$$\alpha_{i,j} = \frac{\sigma(X_{i,j}^{chrom})}{\sigma(Y_{i,j}^{chrom})} \quad (5)$$

$$S_{i,j}^{chrom} = X_{i,j}^{chrom} - \alpha_{i,j} Y_{i,j}^{chrom}$$

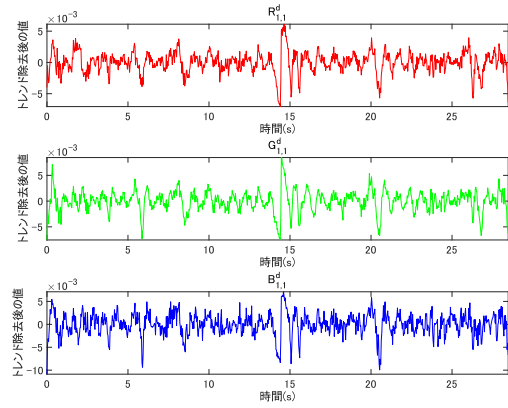


図 3.トレンド除去をした 1 行 1 列目のブロックの各色成分の信号

### Algorithm 3 除外する信号の決定

---

```

 $n \leftarrow$  ブロック数 (クロミナンス信号の数)
 $plist \leftarrow$  長さが  $n$  で各要素が 0 のリスト
 $cnts \leftarrow$  長さが  $n$  で各要素が 0 のリスト
 $fps \leftarrow$  フレームレート
for  $i = 1$  to  $n$  do
     $m \leftarrow S_{i,j}^{chrom}$  の  $i$  列目の信号のピーク間隔平均値
     $plist[i] = 60 \times fps / m$ 
end for
for  $i = 1$  to  $n$  do
     $dlist \leftarrow$   $plist$  の各値を  $plist[i]$  で引いた値の絶対値のリスト
     $cnts[i] = (dlist \leq \text{閾値})$  を満たす個数
end for
 $outliner \leftarrow \text{argmin}(cnts)$ 

```

---

そして, 得られた  $S_{i,j}^{chrom}$  を水平に連結した  $N$  行 12 列の行列を  $S^{chrom}$  とする. この時, 連結する順番としては  $S_{1,1}^{chrom}, S_{1,2}^{chrom}, \dots, S_{3,4}^{chrom}$  の順番で行う.

### 3.3. 主要な信号の抽出

各ブロックから得た 12 個のクロミナンス信号は同じ被写体から得た信号であるため, 脈拍数は等しいと仮定し, ピーク間隔の異なる信号の除外と特異値分解によって主要な信号の抽出を行う. まず, 特異値分解を行う前に各信号のピーク間隔の平均値を計算し, 最も大きな外れ値を持つ信号を探索する. 外れ値の信号の決定方法を Algorithm 3 に示す. この方法は RANSAC と似ており, ある 1 点を通る直線を引いた後, 各点からの差を 10 行目で調べ, 11 行目で設定した閾値以下の個数を数え上げる. そして, これを全ての点について行い, 最もカウント数が少なかったデータ点を外れ値としている. これによって得られる  $outliner$  を  $S^{chrom}$  から除いた  $N$  行 11 列の行列を  $S^{chrom'}$  とする.

$S^{chrom'}$  を得た後は特異値分解を用いて主要な信号の抽出を行う.  $S^{chrom'}$  の特異値分解は式 (6) で与えられる.

$$S^{chrom'} = U \Lambda V^T \quad (6)$$

ただし,  $U$ ,  $V$ ,  $\Lambda$  はそれぞれ  $N$  行  $N$  列の左特異行列, 11 行 11 列の右特異行列, 特異値  $\lambda$  を対角成分に降順で並べた  $N$  行 11 列の行列である.  $\Lambda$  を求めた後は, 特異

値の累積寄与率が  $\alpha$  を超えるとなるような最小の個数, すなわち, 式 (7) を満たす  $k$  のうち, 最小の値を求める.

$$\frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^{11} \lambda_i} \geq \alpha \quad (7)$$

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{11} \geq 0$$

ただし,  $\lambda_i = \Lambda_{i,i}$  である. そして, 式 (6) で計算された  $V$  を  $V_k$  という 11 行  $k$  列の行列にし,  $S_k^{chrom} = S_k^{chrom'} V_k$  という変換を行い,  $N$  行  $k$  列の行列へ変換する.

### 3.4. 脈拍数の決定

$S_k^{chrom}$  を得た後はウェーブレット変換を用いて脈拍数の決定を行う. その前提として, 脈拍に対応する周波数は急激には変わらず長期にわたって観測されるとする. まず, 長期にわたって観測されたということを表現するために  $S_k^{chrom}$  の  $i$  列目 ( $1 \leq i \leq k$ ) の信号を  $x_i(t)$  とし, この信号のウェーブレット変換を考える.  $x_i(t)$  の連続ウェーブレット変換は式 (8) で与えられる.

$$W(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x_i(t) \Psi\left(\frac{t-b}{a}\right) dt \quad (8)$$

ここで,  $a, b, \Psi(\cdot)$  はそれぞれ周波数に相当するスケールパラメータ, 時間に相当するシフトパラメータ, マザーウェーブレットを表す. ここで, 長期にわたって観測された周波数ということを表すために  $\|W(a, b)\|$  を  $b$  で積分し, 振幅スカログラムの時間方向の積分値を計算する. 積分した値  $F_i(a)$  は式 (9) で表される.

$$F_i(a) = \int_{-\infty}^{\infty} \frac{\|W(a, b)\|}{c(b)} db \quad (9)$$

ここで,  $\|\cdot\|$  はノルムを表し,  $c(b)$  は正規化のための値で式 (10) を満たす.

$$c(b) = \max_{0 \leq a \leq fps/2} \|W(a, b)\| \quad (10)$$

ただし,  $fps$  は撮影時のフレームレートを表す. これによって, 式 (9) の被積分関数はある一時刻で見た時, 最大値が 1 になるように正規化, つまり式 (11) を満たすようになる.

$$\forall_b \max_{0 \leq a \leq fps/2} \left( \frac{\|W(a, b)\|}{c(b)} \right) = 1 \quad (11)$$

図 4 は  $S_k^{chrom}$  の 1 列目の信号の  $\|W(a, b)\|/c(b)$  を表し, 横軸が時間, 縦軸が周波数であり, 明るい色ほど値が大きいことを表している. また, 図 5 は式 (9) の計算結果を表し, 横軸が周波数, 縦軸が積分値を表す.

次に, 得られた  $F_i(a)$  だけでは脈拍数を 1bpm 単位で推定できないため,  $x_i$  をフーリエ変換する. この際, 周波数分解能を上げるために信号の長さが  $60 \times fps$  となるように信号の末尾にゼロを付ける. そして,  $x_i$  をフーリエ変換をして得られた振幅スペクトル  $X_i$  と式 (9) で得られた  $F_i$  をフィルタとして用いることで脈拍を表す部分の周波数の振幅を強調する. これを行う際の注意として,  $F_i$  の周波数の間隔がウェーブレット変換を計算機上で行っているため離散値となり,  $X_i$  とのアダマール積を直接計算できないことである. そこで, 3 次のスプライン補間を  $F_i$  に対して行う. また, 脈拍数以外の最大値を取らないようにするために, 補間を行った後, 脈拍数として検出する周波数の区間  $[f_{min}, f_{max}]$  ではない部分の  $F_i$  の値は 0 にする. そして, これらの処理を

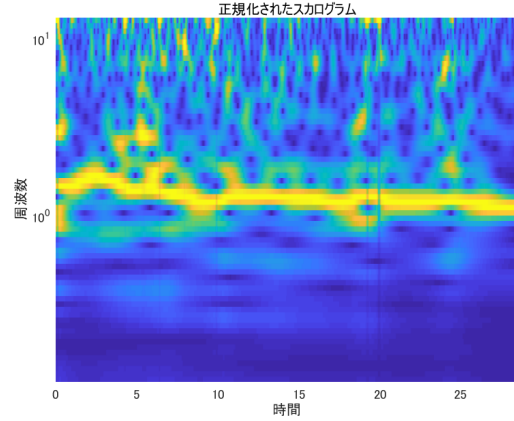


図 4. 正規化されたスカログラム

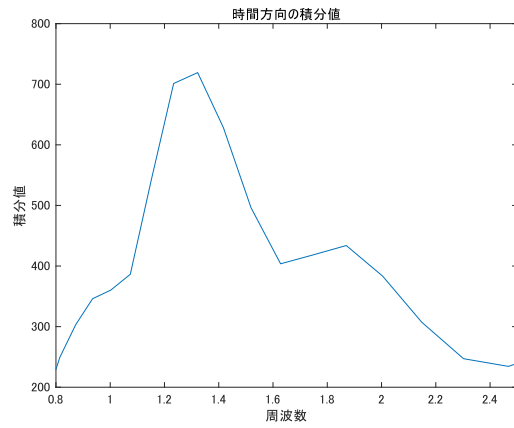


図 5. 図 4 の時間方向の積分値

行った後の  $F_i$  を  $F'_i$  とすると,  $X_i$  に施すフィルタ処理は式 (12) で表される.

$$X'_i(f) = X_i(f) F'_i(f) \quad (12)$$

ここで,  $f$  は周波数を表す. その後,  $X'_i(1 \leq i \leq k)$  の中で最大値が最も大きいスペクトル  $X^*$  を選ぶ. これは式 (13) で表される.

$$X^*(f) := X'_j(f) \quad (13)$$

$$j = \arg \max_{1 \leq i \leq k} (\max_f X'_i(f))$$

最後に,  $f_{min}$  から  $f_{max}$  の周波数帯にあり  $X^*$  の値が最大値を取る周波数  $f^*$  を選ぶ. これは式 (14) で表される.

$$f^* = \arg \max_{f_{min} \leq f \leq f_{max}} X^*(f) \quad (14)$$

そして, 得られた  $f^*$  に 60 を掛け単位を Hz から bpm にし, この値の整数部分を脈拍数の推定結果とする.

## 4. 実験

### 4.1. 実験内容と環境

実験として次の推定方法を比較する.

- 緑色信号のみを用いる方法 [1]



- ICA を用いる方法 [2]
- Chrom 法 [3]
- POS 法 [4]
- 2SR 法 [5]
- 本研究の方法

比較する値としては血圧計で取得した脈拍数の実測値との誤差の平均 MAE, 平均絶対誤差率 MAPE, 実測値との差の標準偏差 SD, 相関係数 corr を用いる. それぞれの値は式 (15), (16), (17), (18) で計算される.

$$MAE = \frac{1}{n} \sum_{i=1}^n |e_i| \quad (15)$$

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{e_i}{ref_i} \right| \quad (16)$$

$$SD = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (e_i - \bar{e})^2} \quad (17)$$

$$corr = \frac{\sum_{i=1}^n estim'_i ref'_i}{\sqrt{\sum_{i=1}^n (estim'_i)^2 \sum_{i=1}^n (ref'_i)^2}} \quad (18)$$

ただし,  $e_i = ref_i - estim_i$ ,  $\bar{e} = \frac{1}{n} \sum_{i=1}^n e_i$ ,  $estim'_i = estim_i - estim$ ,  $ref'_i = ref_i - ref$  であり,  $n$ ,  $estim_i$ ,  $estim$ ,  $ref_i$ ,  $ref$ , はそれぞれデータ数,  $i$  番目の推定値, 推定値の平均値,  $i$  番目の実測値, 実測値の平均値を表す.

実験では iPhone8 をカメラとして利用し, オムロンデジタル自動血圧計 HEM-711 ファジィで取得した脈拍数を実測値に用いる. 撮影条件として, カメラのフレームレートは 30, 解像度は 1080p, カメラは三脚に固定しカメラと顔までの距離は 60cm, 白色蛍光灯を照明に使用している部屋で頭を静止した状態の動画を 105 本撮影した. また, 撮影の開始は血圧計の測定ボタンを押した時, 撮影の終了は血圧計の測定が終わった時とした.

実装と実験は MATLAB R2019a で行い, Image Processing ToolBox, Computer Vision ToolBox, Signal Processing ToolBox, Wavelet ToolBox を用いた. 実験に用いたパラメータやフィルタサイズなどの設定として, トレンド除去で用いる移動平均線を求めるためのフィルタサイズは 30, 特異値の累積寄与率を求める際に使用した式 (7) の  $\alpha$  を 0.9, 脈拍数を検出する帯域  $[f_{min}, f_{max}]$  を  $[0.8, 4]$  すなわち, 脈拍数に換算すると 48bpm から 240bpm とした. また, Algorithm3 の閾値は 5 とした. そして, 緑色成分を用いる方法などバンドパスフィルタが必要な場合は, 0.8Hz から 4Hz までの帯域を通す 6 次のバターワースフィルタを用いた. マザーウェーブレットとして Morse ウェーブレット, フーリエ変換を行う際に掛ける窓関数としてハン窓を用い, Morse ウェーブレットのパラメータ  $\beta$ ,  $\gamma$  をそれぞれ 20 と 3 にした.

## 4.2. 実験結果

実験より, 表 1 が得られた. この表は各方法で推定した時の MAE, MAPE, SD, corr を表している. 表 1 の 1 列目は 4.1 節で示した各推定方法を順に表しており, 2 から 5 列目はそれぞれ式 (15), (16), (17), (18) で計算される値を表している. 表 1 を見ると, 本研究の推定方法は MAE が 0.762bpm, MAPE が 1.04%, SD が 0.889bpm, corr が 0.987 となり, 関連研究のいずれの推定方法と比較しても, 実測値との誤差の改善や相関係数を向上することができた. また, 関連研究の中

表 1. 各推定方法の MAE, MAPE, SD, CORR

処理内容	MAE(bpm)	MAPE(%)	SD(bpm)	corr
緑色信号のみ [1]	1.17	1.60	2.11	0.934
ICA[2]	1.34	1.80	3.35	0.839
Chrom 法 [3]	0.914	1.24	1.20	0.978
POS 法 [4]	0.971	1.31	1.30	0.975
2SR 法 [5]	0.962	1.30	1.29	0.975
本研究	0.762	1.04	0.899	0.987

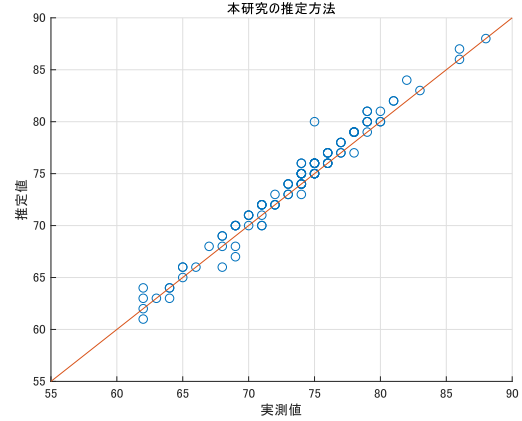


図 6. 本研究の推定結果の散布図

で結果が良かった Chrom 法と比べると, 誤差平均では 0.152bpm, 平均絶対誤差率では 0.20%, 誤差の標準偏差 0.31bpm, 相関係数では 0.009 ほど改善することができた. また, 本研究の推定方法で得られた結果を散布図にすると図 6 のようになった. この図は横軸が実測値, 縦軸が推定値を表しており, 実測値と推定値が一致する部分は赤色の直線が引いてある. また, 縦軸と横軸共に単位は bpm である. 図 6 より, ほとんどのデータについては実測値と近い値を推定でき, 全データの 90% は誤差が 2bpm 未満で収まっていた. しかし, 1 つだけ推定値が実測値より 5bpm 高く推定してしまったデータもあることが分かった.

## 5. 考察

考察として, 本研究と関連研究との違いや分割するブロック数を変えた時の誤差の平均の変化, 本研究の推定方法でも近い値を推定できなかったデータを調べる. 関連研究では酸素化ヘモグロビンの吸光特性を利用した方法 [1] や ICA を用いた方法 [2], Chrom 法 [3] などがあつた. これらの方法はいずれも脈拍数の決定にはフーリエ変換を用いており, 振幅スペクトルが最も大きい値を取る周波数から推定していた. 一方, 関連研究と本研究の大きな違いは, 脈拍数に対応する周波数は急激には変化せず長期にわたって観測されることに着目した部分である. これを脈拍数推定で利用するために本研究では, クロミナンス信号にウェーブレット変換を適用し, スカログラムの時間方向の積分値を計算した. そして, 積分した結果をフィルタとして用い, 振幅スペクトルに掛け, 脈拍数の決定を行った. この結果, Chrom 法などの関連研究の方法よりも精度良く推定でき, いずれの評価指標でも改善することができた.

次にブロック数を変えた時について調べる. 図 7 はそれぞれ行と列方向の分割数を 1 から 9 まで変化した

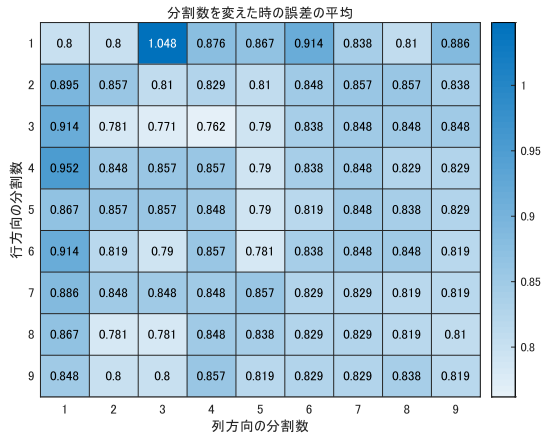


図 7. 分割数を変えた時の誤差の平均

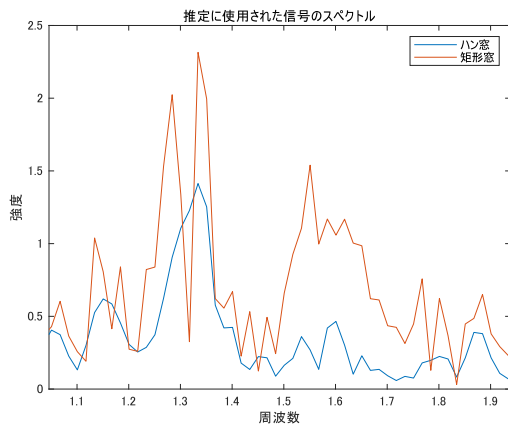


図 8. 推定に使用された信号のスペクトル

時の誤差の平均を表している。また、横軸が列方向の分割数、縦軸が行方向の分割数になっており、色が濃いほど値が大きくなっている。図 7 より、行方向に 3 分割、列方向に 4 分割した場合が最も誤差の平均が小さいことが分かった。また、単に分割数を増やしすぎても精度が向上するわけではなく、むしろ悪くなってしまった。これらの理由を考えるために、皮膚領域を 3 行 4 列に分割した時の画像である図 1 中の皮膚領域を見る。分割した皮膚領域の画像より、3 行 4 列に分割したことによって鼻の穴の周辺や頬といった領域に分割されていることが分かる。これによって、脈拍の取りやすい頬の部分と取りにくい鼻の穴の周辺が別々に信号へ変換されたため、ブロックで分けなかった時よりも精度が向上したのだと考えられる。また、分割数を増やしすぎても精度が向上しなかった理由は 1 ブロックあたりの面積が小さくなりすぎたためだと考えられる。そのため信号へ変換する際、ピクセル数が十分に足りず脈拍数以外のノイズが信号に残ってしまったため精度が向上しなかったのだと考えられる。

最後に、本研究の推定方法で近い値を推定できなかったデータ (図 6 で実測値が 75bpm、推定値が 80bpm のデータ) について調べる。このデータで推定に用いられた信号のハン窓を掛けた時の振幅スペクトルと矩形窓を掛けた時の振幅スペクトルは図 8 のようになった。この図の横軸は周波数、縦軸は強度を表しており、青の線がハン窓を掛けた時、赤の線が矩形窓の時を表している。図 8 より、ハン窓を掛けた時には現れなかった脈拍数

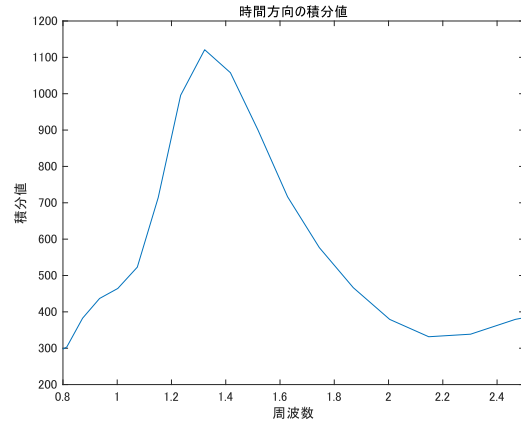


図 9. スカログラムの時間方向の積分値

に対応するピークが矩形窓を掛けた時には現れており、強度は低いものの全く観測されなかったというわけではなかった。一方、この信号のスカログラムの時間方向積分値は図 9 のようになった。この図は横軸が周波数、縦軸が積分値を表している。図 9 を見ると、図 8 とは違って脈拍数付近にはピークは無くなっていることが分かった。そのため、式 (12) をどちらの振幅スペクトルに行っても 75bpm に近い値を推定することはできなかった。このことより、脈拍のピークと別の強いピークとの間隔が近いと正しく推定できなくなると考えられる。

## 6. まとめ

本研究では皮膚のどの箇所を調べても脈拍数は同じであることと、脈拍数に対応する周波数は急激には変化せず長期にわたって観測されることに着目し、主要な信号の抽出とウェーブレット変換を用いた脈拍数推定方法について述べた。実験の結果、推定値と実測値との誤差の平均は 0.762bpm、平均絶対誤差率は 1.04%、差の標準偏差は 0.899bpm、相関係数は 0.987 となり、全データの 90% は誤差が 2bpm 未満で収まった。今後の課題として、推定方法の改良により推定精度をより高くすることや、スマートフォンを手で持った状態や振動がある環境下など実生活の状況に近い撮影条件で推定した時に高精度で推定できるようにすることであると考えられる。

## 参考文献

- [1] W. Verkruijsse, L.O. Svaasand, and J.S. Nelson, "Remote plethysmographic imaging using ambient light", *Optics express*, vol.16, no.26, pp.21434-21445, Dec. 2008.
- [2] D. McDuff, S. Gontarek, and R.W. Picard, "Improvements in Remote Cardiopulmonary Measurement Using a Five Band Digital Camera", *IEEE Transactions on Biomedical Engineering*, vol.61, no.10, pp.2593-2601, Oct. 2014.
- [3] G de Haan., and V Jeanne, "Robust Pulse Rate From Chrominance-Based rPPG", *IEEE Transactions on Biomedical Engineering*, vol.60, no.10, pp.2878-2886, Oct. 2013.
- [4] W. Wang, A. C. den Brinker, S. Stuijk and G. de Haan, "Algorithmic Principles of Remote PPG," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 7, pp. 1479-1491, July 2017.
- [5] W. Wang, S. Stuijk and G. de Haan, "A Novel Algorithm for Remote Photoplethysmography: Spatial Subspace Rotation," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 9, pp. 1974-1984, Sept. 2016.
- [6] 坂巻燎, 藤田悟, "主要信号抽出とウェーブレット変換による非接触脈拍数推定法", 情報処理学会第 83 回全国大会, 2021.